

Tight Bounds for Broadcasting in the Linear Cost Model

Bruno Beauquier — Olivier Delmas — Stéphane Pérennes

N° 3827

Décembre 1999

THÈME 1



***Rapport
de recherche***

Tight Bounds for Broadcasting in the Linear Cost Model

Bruno Beauquier ^{*}, Olivier Delmas [†], Stéphane Pérennes ^{*}

Thème 1 — Réseaux et systèmes

Projet Sloop

Rapport de recherche n° 3827 — Décembre 1999 — 21 pages

Abstract: This work considers broadcast protocols made of successive communication rounds in the *linear cost* model: the time needed to send a message of length L is defined as $\alpha + L\tau$. In this model, the *communication time* of any algorithm \mathcal{A} is expressed as the sum $R_{\mathcal{A}} \cdot \alpha + T_{\mathcal{A}} \cdot \tau$, where $R_{\mathcal{A}}$ is the *number of rounds* and $T_{\mathcal{A}}$ the *transmission cost* of the algorithm. In order to design an efficient algorithm realizing a given communication pattern, it appears that minimizing $R_{\mathcal{A}}$ and $T_{\mathcal{A}}$ are antinomic goals. We study this trade-off issue for broadcast protocols. Surprisingly, such a general theoretical study has almost never been done. In the literature, only the two opposite issues are actually considered: minimizing the number of rounds in the case of short messages, or minimizing the transmission cost in the case of large messages. Our results concern the fully-connected N -nodes network K_N , with $N = (k + 1)^T$, in the bidirectional k -ports mode. We derive tight bounds on the communication time for broadcasting in $T + r$ rounds, our lower bounds holding for any network.

Key-words: communication networks, broadcasting, linear cost.

An extended abstract of this paper written by the second author has been accepted in the *5th International Colloquium on Structural Information and Communication Complexity (SIROCCO'98)*.

This work has been partially supported by the AFIRST in the framework of the French-Israeli project *Communication Algorithms in Optical Networks* and by the French RNRT project *PORTO*.

^{*} SLOOP is a joint project with the CNRS and the University of Nice-Sophia Antipolis (I3S laboratory). E-mail: {Bruno.Beauquier,Stephane.Perennes}@inria.fr.

[†] LaBri, Université Bordeaux 1, 351 cours de la Libération, 33405 Talence Cedex, France. E-mail: Olivier.Delmas@labri.u-bordeaux.fr.

Bornes Fines pour la Diffusion avec Coût Linéaire

Résumé : Ce travail considère des protocoles de diffusion se déroulant en *étapes* successives sous le modèle *linéaire* du coût de communication : le temps nécessaire pour envoyer un message de longueur L est défini par $\alpha + L\tau$. Dans ce modèle, le *temps de communication* d'un algorithme s'exprime par la somme $R_A \cdot \alpha + T_A \cdot \tau$, où R_A est le *nombre d'étapes* et T_A le *coût de transmission* de l'algorithme. En vue de concevoir un algorithme efficace réalisant un schéma de communication donné, il apparaît que minimiser R_A et T_A sont des objectifs antinomiques. Nous étudions le compromis nécessaire entre ces deux coûts dans le cadre de la diffusion. Il s'avère qu'une telle étude théorique générale n'a pratiquement jamais été entreprise. Dans la littérature, seulement les deux cas extrêmes sont considérés : minimisation du nombre d'étapes dans le cas de messages courts, ou minimisation du coût de transmission dans le cas de messages longs. Nos résultats concernent le réseau complet K_N , avec $N = (k+1)^T$, sous le mode k -ports bidirectionnel. Nous dérivons des bornes fines sur le temps de communication pour la diffusion en $T + r$ étapes, les bornes inférieures étant valables dans tout réseau.

Mots-clés : réseaux de communication, diffusion, coût linéaire.

1 Introduction

When studying information dissemination in communication networks, often considered are protocols made of successive communication *rounds* with a *linear* model for the communication *cost*: the time needed to send a message of size L from a node to another is defined as $\alpha + L\tau$, where α stands for a start-up time while $L\tau$ represents the cost of sending L bits of data in a channel with bandwidth $1/\tau$. The *cost of a round* is defined as the maximum of the cost of all the communications done during this round. In such a modelization, the **communication time** of any algorithm \mathcal{A} can be expressed as the sum $R_{\mathcal{A}} \cdot \alpha + T_{\mathcal{A}} \cdot \tau$, where $R_{\mathcal{A}}$ is the **number of rounds** and $T_{\mathcal{A}}$ the **transmission cost** of the algorithm. This linear model of the communication cost is often used in the literature and references to algorithmic work can be found in [6, 7]. When trying to design an optimal algorithm performing a given communication pattern, it appears that minimizing $R_{\mathcal{A}}$ and $T_{\mathcal{A}}$ are often *antinomic* goals. In fact, in the literature many authors optimize only one of these two parameters by considering one of the two extremal cases:

- *Short messages (i.e. $\alpha \gg L\tau$)* – In this case, $R_{\mathcal{A}} \cdot \alpha$ is the leading term of the communication time. Hence, only the number of rounds is to be minimized. Many authors have proposed broadcast algorithms optimal with that respect (see the surveys [10, 11], the book [14], the chapter [13] or the theses [6, 7]).
- *Large messages (i.e. $L\tau \gg \alpha$)* – In this case, only the transmission cost is to be minimized. Techniques, like pipelined algorithms using “disjoint” spanning trees, ensure a nearly optimal cost for message transmission (see [8, 12, 15] for hypercube network or [4, 9] for grids, and [10, 14] for a general survey). Note that this kind of optimality always implies a large number of communication rounds.

Some authors considered also families of algorithms $\mathcal{A}(p)$ depending on some parameter p (for example, for pipelined algorithms p usually represents the size of packets, assuming that all packets have the same size – see the survey [10] and the associated references). In this case, the family $\mathcal{A}(p)$ uses $R_{\mathcal{A}}(p)$ rounds with a transmission cost $T_{\mathcal{A}}(p)$. Within a family, as the functions $R_{\mathcal{A}}(p)$ and $T_{\mathcal{A}}(p)$ have usually an opposite behavior, the optimal p minimizing the communication time can be chosen for given α , L and τ .

Here, in order to design optimal algorithms and to study this trade-off, we proceed in a different way. Given the number of rounds, we try to determine (at least to find a tight estimation of) the minimum transmission cost of any algorithm realizing the communication pattern within this given number of rounds. A general theoretical study about such a trade-off has almost never been done (except in [5]) and therefore no optimal protocol was known but in the case of short or large messages, by optimizing only the leading term.

The determination of the minimum transmission cost of an algorithm using a given number of rounds appears to be a difficult problem. So, we restrict ourselves to the communication scheme called broadcasting, where a single source has to send the same message of length L to all the other nodes in the network. We also assume that links are full-duplex (*i.e.* there is a bidirectional link between each pair of node) and that a node can simultaneously send and receive data on at most k -ports. We consider the case where the network is the complete graph K_N and we assume that $N = (k + 1)^T$, as this is the maximum number of nodes which can be informed in T rounds. With these hypotheses, any broadcast algorithm needs *at least* T rounds.

In this paper, we study the antinomy between the number of rounds and the transmission cost. When the number of rounds is small, more exactly of the form $T + r$ with $r \leq \log_{k+1}(T) + 1$, we show that an estimation of the transmission cost is given by an exponential decrease $\frac{T}{(k+1)^r} + \Theta(\frac{1}{k})$. Conversely, for a large number of extra rounds $r \geq T - 1$, we observe a linear decrease $\frac{T+r}{kr+1} - \mathcal{O}(\frac{1}{r^2})$, whereas for intermediate values no trade-off occurs, the transmission cost remaining close to $2/k$.

The article is organized as follows. In section 2 we give the definitions and hypotheses and show why our study is valid for both the store-and-forward and the circuit-switched routing modes. We also introduce a function allowing to study the trade-off issue formally. In section 3 we give an upper bound for broadcasting with a few extra rounds, while we derive tight lower bounds in section 4. Other tight bounds for many extra rounds are derived in section 5. In conclusion we summarize our results and discuss their applications.

2 Framework

Recall that we consider *synchronous* protocols made of successive communication rounds.

Definition 1 The **communication time** of a broadcast algorithm \mathcal{A} in a network G can be expressed as $b(\mathcal{A}) = R_{\mathcal{A}} \cdot \alpha + T_{\mathcal{A}} \cdot \tau$, where $R_{\mathcal{A}}$ is the number of rounds and $T_{\mathcal{A}}$ the transmission cost. The minimum broadcast time over all algorithms is denoted by $b(G)$.

For any k -ports network with N nodes, the usual lower bound on the time for broadcasting a message of length L is expressed as the sum of the lower bounds valid for the number of rounds and for the transmission cost [6]:

$$b(G) \geq \lceil \log_{k+1} N \rceil \cdot \alpha + \frac{L}{k} \cdot \tau \quad (1)$$

As already noticed, existing upper bounds have quite a different form where a trade-off appears.

The function defined below will be useful to derive a relation between the *number of rounds* and the *transmission cost*.

Definition 2 Let G be a fixed network and T the minimum number of rounds for broadcasting in G . Then, $f(G, L, r)$ is defined as the **minimum transmission cost** of any algorithm broadcasting in G a message of length L and completing in r **extra rounds**, *i.e.* using $T + r$ rounds.

With the previous definition, we have $b(G) \geq (T + r) \cdot \alpha + f(G, L, r) \cdot \tau$, for any $r \geq 0$.

Hypothesis 3 Packet sizes are assumed to be **continuous**.

This does not hold formally in practice, as messages are made of integer numbers of bits. However, as these numbers are generally large, the continuity assumption makes sense.

Property 4 Assuming continuous packet sizes, the function f is linear in L , *i.e.* $f(G, L, r) = f(G, 1, r) \cdot L$. Moreover, $f(G, 1, r)$ is a non-increasing function of r .

Proof. Let M be a message of length L and let $\mathcal{A}_{L'}$ be an algorithm broadcasting a message M' of length L' with transmission cost $f(G, L', r)$.

Let us denote by h the function $x \mapsto x \cdot L/L'$ from $[0, L']$ to $[0, L]$. Note that h induces a function H from the set of submessages of M' to the set of submessages of M . With function H we are able to define submessages of M from submessages of M' . By following the same communication protocol as $\mathcal{A}_{L'}$ with the images by H , we obtain an algorithm \mathcal{A}_L broadcasting M with cost $f(G, L', r) \cdot L/L'$. Hence, $f(G, L, r) \leq f(G, L', r) \cdot L/L'$. By inverting roles of M' and M , we get the reverse inequality.

Note at last that any algorithm using $r + 1$ rounds can emulate an algorithm using r rounds by sending empty messages during its last round, so $f(G, 1, r + 1) \leq f(G, 1, r)$. \square

Remark 5 As $f(G, L, r)$ is linear in L , we will consider w.l.o.g. that $L = 1$.

Hypothesis 6 We consider the case of the **complete graph** K_N with N nodes.

Note that the behaviour of $f(G, 1, r)$ strongly depends on the topology of network G . Any link conflict is avoided in the complete graph. Therefore we will derive lower bounds valid for every network with N nodes. Moreover, as each pair of vertices can communicate

in the complete graph, the study is also valid for both the store-and-forward and the circuit-switched routing modes.

Finally, in order to simplify computations we will assume that:

Hypothesis 7 Under the k -ports communication mode, $N = (k + 1)^T$.

This is not a very restrictive hypothesis, as N is thereby the maximum number of nodes which can be informed by a broadcast protocol completing in T communication rounds.

With all these hypotheses, we give in the next sections results on $f(K_N, 1, r)$, with $N = (k + 1)^T$, that will be denoted simply by $f_T(r)$.

With this notation, we have immediately:

Proposition 8 $f_T(0) = T$.

Proof. As $N = (k + 1)^T$, every node other than the source can be informed only once during the T rounds and must receive the whole message whose length is 1. Hence, the size of any message sent during any round must be equal to 1, and so $f_T(0) \geq 0$. A greedy algorithm matching this bound can be defined as follows. During T rounds every informed node sends the whole message to k not yet informed nodes. The broadcast is thus completed and the resulting transmission cost is T . \square

3 Upper bound for broadcasting with a few extra rounds

Here, we construct inductively algorithms for broadcasting with r extra rounds. We initiate the induction by using the greedy broadcast algorithm given in Proposition 8.

Lemma 9 For $T > 0$ and $r > 0$, $f_T(r) \leq \frac{2}{k+1} + \frac{f_{T-1}(r-1)}{k+1}$.

Proof. For $t \geq 0$, let $N_t = (k + 1)^t$, hence $N = N_T$. The idea consists in defining an algorithm in K_N using $T + r$ rounds from an algorithm in $K_{N_{T-1}}$ using $T - 1 + r - 1$ rounds.

The vertex set $V(K_N)$ of the complete graph K_N can be seen as the cartesian product $\{1, 2, \dots, k + 1\} \times \{1, 2, \dots, N_{T-1}\}$ and can be represented by the following matrix with $k + 1$

rows and N_{T-1} columns:

$$V(K_N) = \begin{pmatrix} (1,1) & (1,2) & \cdots & (1,N_{T-1}) \\ (2,1) & (2,2) & \cdots & (2,N_{T-1}) \\ \vdots & \vdots & \ddots & \vdots \\ (k+1,1) & (k+1,2) & \cdots & (k+1,N_{T-1}) \end{pmatrix}$$

The source $(1,1)$ first partitions equally the message into $k+1$ packets each of size $1/(k+1)$, denoted by m_i for $1 \leq i \leq k+1$.

- During the first round, the source $(1,1)$ sends packet m_i to node $(i,1)$, for $2 \leq i \leq k+1$. The transmission cost of this round is $1/(k+1)$.
- During the next $(T-1+r-1)$ rounds and for $1 \leq i \leq k+1$, each packet m_i is broadcast from node $(i,1)$ to all the vertices in line i . This is realized by performing $k+1$ independent and parallel identical protocols broadcasting in $(T-1)+(r-1)$ rounds a message of length $1/(k+1)$ in $k+1$ disjoint complete graphs of cardinality N_{T-1} . By definition of f_{T-1} , the transmission cost of this phase is $f_{T-1}(r-1)/(k+1)$.
- Finally, as the original message is distributed on every column, each node (i,j) sends in the last round packet m_i to the k others nodes in column j (parallel total exchanges of the packets m_i are realized in the different columns). The transmission cost of this last round is $1/(k+1)$.

Summarizing, we obtain the inductive inequality $f_T(r) \leq \frac{2}{k+1} + \frac{f_{T-1}(r-1)}{k+1}$. \square

Now the following proposition can be stated.

Proposition 10 For $r \leq T$, $f_T(r) \leq \frac{T-r}{(k+1)^r} + \frac{2}{k} \left(1 - \frac{1}{(k+1)^r}\right)$.

Proof. For $r = 0$ we use Proposition 8. For $1 \leq r \leq T$, applying inductively Lemma 9 leads to $f_T(r) \leq \sum_{i=1}^r \frac{2}{(k+1)^i} + \frac{f_{T-r}(0)}{(k+1)^r}$. Note that the condition $r \leq T$ follows from the fact that Lemma 9 can be applied only for $T > 0$. \square

4 Lower bounds for broadcasting with a few extra rounds

4.1 Notation

For every $t \geq 0$, let $N_t = (k+1)^t$. Then N_t is an upper bound on the maximum number of nodes having received some information after the t first rounds. For $1 \leq t \leq T+r$, we need the following definitions:

$$\begin{aligned} l_t & : \text{maximum size of the messages sent during round } t, \\ \mathcal{L}_t = \sum_{i=1}^t l_i & : \text{transmission cost after the first } t \text{ rounds,} \\ i_t & : \text{total information received during the round } t, \\ \mathcal{I}_t = \sum_{i=1}^t i_i & : \text{total information received after the first } t \text{ rounds,} \end{aligned}$$

and by convention, $\mathcal{L}_0 = \mathcal{I}_0 = 0$.

4.2 Preliminaries

To study the transmission cost once some vertices have already been informed, we need to consider particular protocols where a given number s of sources can simultaneously broadcast information, and which we call *broadcast protocols with s sources*.

Lemma 11 *For any broadcast protocol with s sources, for any $t \geq 1$ and for a fixed \mathcal{L}_t , we have $\mathcal{I}_t \leq s \cdot \frac{\mathcal{L}_t}{t} \cdot (N_t - 1)$. Moreover, equality holds only if all the costs l_i , with $1 \leq i \leq t$, are equal.*

Proof. The following holds for each round $t+1$ such that $t \geq 0$.

- At most sN_t nodes can send messages, hence $i_{t+1} \leq k(sN_t)l_{t+1}$.
- Let $i(x, t)$ be the total information received by node x after t rounds. The information sent by the sources is at most $s(kl_{t+1})$ and each of the other nodes can send no more than what it has already received, hence $i_{t+1} \leq s(kl_{t+1}) + k \sum_{x \in G} i(x, t) = skl_{t+1} + k\mathcal{I}_t$.

So for every t such that $0 \leq t \leq T$, we have:

$$\mathcal{I}_{t+1} \leq \mathcal{I}_t + ksN_t l_{t+1} \tag{2}$$

$$\mathcal{I}_{t+1} \leq (k+1)\mathcal{I}_t + ksl_{t+1} \tag{3}$$

Note that the lemma is true for $t = 1$, as we have $N_0 = 1$, $\mathcal{I}_0 = 0$, leading to $\mathcal{I}_1 \leq ksl_1 = sl_1(N_1 - 1)$. We assume that it holds for a given t and we prove it for $t+1$.

We recall that \mathcal{L}_{t+1} is fixed. For a protocol with some transmission cost l_{t+1} at round $t + 1$, we have $\mathcal{L}_t = \mathcal{L}_{t+1} - l_{t+1}$ and by induction hypothesis

$$\mathcal{I}_t \leq s \cdot \frac{\mathcal{L}_{t+1} - l_{t+1}}{t} \cdot (N_t - 1)$$

From (2) and (3), we get:

$$\begin{aligned} \mathcal{I}_{t+1} &\leq \mathcal{I}_t + ksN_t l_{t+1} \\ &\leq s \left(\frac{\mathcal{L}_{t+1} - l_{t+1}}{t} (N_t - 1) + kN_t l_{t+1} \right) = \varphi_1(l_{t+1}) \\ \mathcal{I}_{t+1} &\leq (k+1)\mathcal{I}_t + ks l_{t+1} \\ &\leq s \left(\frac{\mathcal{L}_{t+1} - l_{t+1}}{t} (N_t - 1)(k+1) + k l_{t+1} \right) = \varphi_2(l_{t+1}) \end{aligned}$$

Functions φ_1 and φ_2 are two affine functions in l_{t+1} , φ_1 being increasing while φ_2 is non-increasing. Therefore, $\min(\varphi_1(l_{t+1}), \varphi_2(l_{t+1}))$ is smaller than the common value obtained by solving $\varphi_1(l_{t+1}) = \varphi_2(l_{t+1})$. The solution of this equation is $l_{t+1} = \frac{\mathcal{L}_{t+1}}{t+1}$ and we obtain:

$$\mathcal{I}_{t+1} \leq \varphi_1(l_{t+1}) = \varphi_2(l_{t+1}) = \frac{\mathcal{L}_{t+1}}{t+1} \cdot (N_{t+1} - 1)$$

To reach this bound, we must have by induction $l_{i+1} = \frac{\mathcal{L}_{i+1}}{i+1}$ for each $i \leq t$, which indeed implies that all the intermediate costs are equal. In this view, the transmission optimization may be called *greedy*. \square

Corollary 12 For $0 \leq p < q \leq T + 1$, the information transmission \mathcal{I}_q after q rounds of every protocol having fixed values of \mathcal{I}_p , \mathcal{L}_p and \mathcal{L}_q satisfies:

$$\mathcal{I}_q \leq \mathcal{I}_p + \frac{\mathcal{L}_q - \mathcal{L}_p}{q - p} \cdot (N_q - N_p) \quad (4)$$

Proof. First, note that after p rounds at most N_p nodes can be informed. So, after round p , an information transmission occurs from N_p sources during $q - p$ rounds and with a fixed cost $(\mathcal{L}_q - \mathcal{L}_p)$. According to Lemma 11 applied with $t = q - p$, $\mathcal{L}_t = \mathcal{L}_q - \mathcal{L}_p$ and $s = N_p$ we have $(\mathcal{I}_q - \mathcal{I}_p) \leq N_p \left(\frac{\mathcal{L}_q - \mathcal{L}_p}{q - p} (N_{q-p} - 1) \right)$ and the result holds. \square

Remark 13 For values of t greater than $T + 1$, Lemma 11 does not give an accurate estimation of \mathcal{I}_t . This is due to the fact that for $t \geq T + 1$, the maximum number of

informed nodes is N_T , which is much smaller than N_t . Nevertheless, lower bounds on $f_T(r)$ can still be derived from inequalities (2) and (3), but they are not tight enough and the refined analyses given in the next sections allow to obtain better results.

Consider a dissemination protocol with one source using $T+r$ rounds. After round $T+r$, each node must have received a complete message of length 1, hence, if $i(x, t)$ (as in the proof of Lemma 11) denotes the total information received by node x after t rounds, we have for any node x , $i(x, T+r) \geq 1$. By summing over all the nodes (except the source) we get the next inequality:

$$\mathcal{I}_{T+r} \geq N_T - 1 \quad (5)$$

Remark 14 Using the fact that any broadcast algorithm using T rounds satisfies $\mathcal{I}_T \geq N - 1$, we find again Proposition 8 from Lemma 11, *i.e.* the transmission cost \mathcal{L}_T is at least T .

4.3 One extra round

Proposition 15 $f_T(1) = \frac{T+1}{k+1}$

Proof. We prove that $f_T(1) \geq \frac{T+1}{k+1}$ by reduction to the absurd, assuming the existence of an algorithm using $T+1$ rounds with a transmission cost $\mathcal{L}_{T+1} < \frac{T+1}{k+1}$.

First we prove that $l_{T+1} < \frac{1}{k}$. With inequalities (5) and (3) and from Lemma 11, we have:

$$\begin{aligned} N_T - 1 &\leq \mathcal{I}_{T+1} \\ &\leq (k+1)\mathcal{I}_T + kl_{T+1} \end{aligned}$$

$$\mathcal{I}_T \leq \frac{\mathcal{L}_T}{T}(N_T - 1)$$

$$\begin{aligned} \text{by hypothesis, } \mathcal{L}_T &= \mathcal{L}_{T+1} - l_{T+1} \\ &< \frac{T+1}{k+1} - l_{T+1} \end{aligned}$$

Combining the three relations leads to $l_{T+1} < \frac{1}{k+1 - \frac{kT}{N-1}} \leq \frac{1}{k}$.

As $kl_{T+1} < 1$, a node without information cannot receive enough information during the last round. So, during the first T rounds each node must receive some information, exactly one message as $N = (k+1)^T$.

Let $l = \min \{l_t \mid 1 \leq t \leq T+1\}$ be the minimum cost of all rounds. We have $(T+1)l \leq \mathcal{L}_{T+1} < \frac{T+1}{k+1}$, so $l < \frac{1}{k+1}$. Now, we distinguish two cases.

- The minimum value l is attained for l_{T+1} . As a vertex can receive at most kl information in round $T+1$, then it should be know at least $1 - kl$ information after round T and so, according to Proposition 8 we have $l_t \geq 1 - kl$ for $1 \leq t \leq T$.
- The minimum value l is attained for some t_0 with $1 \leq t_0 \leq T$. As the algorithm follows a spanning tree pattern during the first T rounds, exactly $(N_{t_0} - N_{t_0-1})$ nodes receive at most l information during round t_0 . Let S be the set of all the descendants of these nodes after round T ; and note that $|S| = (k+1)^{T-t_0}(N_{t_0} - N_{t_0-1}) = N_T - N_{T-1}$. So, all the nodes in S have at most l information after round T . As $l < 1/(k+1)$, no node in S can receive, in the last round, k messages from nodes in S . As $k(N - |S|) = |S|$, we deduce that during round $T+1$ each node in S receives $(k-1)$ messages of length at most l from nodes in S and one message of length at least $(1 - kl)$ from a node not in S . Consequently, any node not in S must have at least $(1 - kl)$ information after round T . So, $l_{T+1} \geq 1 - kl$ and $l_t \geq 1 - kl$ for $1 \leq t \leq T$, with $t \neq t_0$.

In both cases we have $\mathcal{L}_{T+1} \geq l + T(1 - kl) = \varphi(l)$. Finally, just note that $\varphi(l)$ is a strictly decreasing function in l . As $l < \frac{1}{k+1}$, we have $\mathcal{L}_{T+1} > \varphi(\frac{1}{k+1}) = \frac{T+1}{k+1}$, contradicting our main hypothesis. As the conditions used in this proof are necessary to perform any broadcast protocol, then, it is proved that $f_T(1) \geq \mathcal{L}_{T+1} \geq \frac{T+1}{k+1}$ and the result follows from Proposition 10. \square

4.4 Two extra rounds

Before to state a lower bound for two extra rounds, we need to introduce a new notion about the *transmission cost*.

Definition 16 Given a network G with one source vertex. A **transmission algorithm** using t rounds is a communication protocol such that:

- Initially the source has a message of infinite size, and all the other vertices in G have no information.
- During the t rounds of the protocol, each vertex with some information can send k different submessages such that each size is not greater than the known information.
- At the end of the protocol, *i.e.* after t rounds, all the vertices in G have received an amount of information at least equal to 1.

Remark 17 Clearly, these transmission constraints are weaker than those for broadcasting. Indeed, at the end of a transmission protocol, we make no assumption about what message reaches each vertex. We only restrict the final amount of information in each vertex.

Definition 18 Let $g_T(r)$ be the **minimal transmission cost** of any transmission algorithm for the complete graph K_N using $T + r$ rounds.

Property 19 *As any broadcast algorithm satisfies the transmission constraints, we have $f_T(r) \geq g_T(r)$.*

Our lower bound can now be stated.

Proposition 20 $f_T(2) = \frac{T}{(k+1)^2} + \frac{2}{k+1} - \Theta\left(\frac{1}{kT}\right)$.

Proof. From Proposition 10 we have $f_T(2) \leq \frac{T}{(k+1)^2} + \frac{2}{k+1}$. We consider an optimal broadcast algorithm using $T + 2$ rounds.

After the first round, at least $(1 - kl_1)$ information remains to be broadcast from the source in $T + 1$ rounds, so $l_1 + (1 - kl_1)f_T(1) \leq \mathcal{L}_{T+2} \leq \frac{T}{N_2} + \frac{2}{N_1}$ and from Proposition 15 we deduce:

$$l_1 \geq \frac{kT - 1 - k}{(kT - 1)N_1} = \frac{1}{N_1} - \Theta\left(\frac{1}{kT}\right) \quad (6)$$

Before the last round, every node must have at least $(1 - kl_{T+2})$ information. Therefore, by definition of the function g_T , we get that \mathcal{L}_{T+1} is at least $(1 - kl_{T+2})g_T(1)$. So l_{T+2} satisfies the same inequality as l_1 :

$$l_{T+2} \geq \frac{kT - 1 - k}{(kT - 1)N_1} = \frac{1}{N_1} - \Theta\left(\frac{1}{kT}\right) \quad (7)$$

Let λ be such that $\mathcal{L}_{T+2} = l_1 + T\lambda + l_{T+2}$. By applying Corollary 12 with $p = 1$ and $q = T + 1$, we obtain $\mathcal{I}_{T+1} \leq k + \lambda(N_{T+1} - N_1)$. Then from $N_T - 1 \leq \mathcal{I}_{T+2} \leq N_1\mathcal{I}_{T+1} + k$ we deduce:

$$\lambda \geq \frac{1}{N_2} - \Theta\left(\frac{1}{N}\right) \quad (8)$$

As T/N is small compared to $1/(kT)$, combining (6), (7) and (8) gives:

$$f_T(2) = \mathcal{L}_{T+2} = l_1 + T\lambda + l_{T+2} \geq \frac{T}{N_2} + \frac{2}{N_1} - \Theta\left(\frac{1}{kT}\right)$$

□

4.5 Some extra rounds

We have given in section 3 algorithms for broadcasting with r extra rounds, inducing upper bounds on $f_T(r)$. For large enough values of T , their transmission cost can be approximated by the sum of the two terms $\frac{T}{(k+1)^r}$ and $\frac{2}{k}$. The former is dominating for small values of r and decreases exponentially with r . We prove in the following that in this case the upper bounds of Proposition 10 are tight. Therefore we show the efficiency of our algorithms, which allow to decrease significantly the transmission cost with a few additive rounds.

Theorem 21 For $0 \leq r \leq T$, $f_T(r) = \frac{T}{(k+1)^r} + \Theta\left(\frac{1}{k}\right)$.

Proof. We derive here a lower bound on $f_T(r)$ matching the upper bound given by Proposition 10. Let consider a broadcast algorithm using $T + r$ rounds. After round $r - 1$ it can be seen as a broadcast protocol with (at most) N_{r-1} sources. So from inequality (3) with $t = T + r - 1$ we get:

$$\mathcal{I}_{T+r} \leq (k+1) \cdot \mathcal{I}_{T+r-1} + k \cdot N_{r-1} \cdot l_{T+r}$$

By iteration for $T \leq t < T + r - 1$ and as $l_{t+1} \geq 1$, we obtain:

$$\mathcal{I}_{T+r} \leq N_r \cdot \mathcal{I}_T + \sum_{i=0}^{r-1} (k \cdot N_i \cdot l_{T+r-i}) \leq N_r \cdot \mathcal{I}_T + (N_r - 1)$$

Then, we apply Lemma 11 with $t = T$,

$$N_T \leq \mathcal{I}_{T+r} + 1 \leq N_r \cdot (\mathcal{I}_T + 1) \leq N_r \cdot \left(\frac{\mathcal{L}_T}{T} (N - 1) + 1 \right)$$

and it follows that $\mathcal{L}_T \geq T \cdot \frac{N_{T+r-1}-1}{N-1} = \frac{T}{N_r} + \frac{T}{N-1} \cdot \frac{1-N_r}{N_r} = \frac{T}{N_r} + \Theta\left(\frac{T}{N}\right)$.

Now, we must compare this relation with the upper bound given by Proposition 9, that is we compare $\frac{T}{N}$ with $\frac{1}{k}$. As $\frac{T}{N}$ is small compared to $\frac{1}{k}$ the proposition holds. \square

The above proposition is meaningless when $r > \log_{k+1}(T)$ as it reduces to $f_T(r) = \Theta\left(\frac{1}{k}\right)$. In the following we investigate this case and we give a better bound.

Proposition 22 For $r < T - 1$, $f_T(r) \geq \frac{2}{k}$.

Proof. We consider a broadcast algorithm using $T + r$ rounds and having a transmission cost strictly less than $2/k$, and we show that $r \geq T - 1$. Let t_0 be the last round after which

it remains some information not yet broadcast from the source, hence $\mathcal{L}_{t_0+1} \geq 1/k$. After round t_0 , T rounds are necessary to broadcast this last piece of information in the entire network, so $t_0 + T \leq T + r$ and $t_0 \leq r$. Given the assumption on the total transmission cost, every node must have received some information after the round $t_0 + 1$, otherwise $\mathcal{L}_{T+r} \geq \mathcal{L}_{t_0+1} + 1/k \geq 2/k$. Therefore, we have $t_0 + 1 \geq T$, leading to $r \geq T - 1$. \square

Corollary 23 For $\log_{k+1}(T) + 1 < r < T - 1$, $f_T(r) = \frac{2}{k} + \mathcal{O}\left(\frac{1}{k^2}\right)$.

Proof. By Proposition 10 we have $f_T(r) \leq \frac{2}{k} + \frac{k(T-r)-2}{k(k+1)^r}$. As for $r \geq \log_{k+1}(T) + 2$, we have $(k+1)^r \geq T \cdot (k+1)^2$, then, $f_T(r) \leq \frac{2}{k} + \mathcal{O}\left(\frac{1}{k^2}\right)$ and the result holds with Proposition 22. \square

Remark 24 It has been proved that the minimum transmission cost $f_T(r)$ decreases exponentially when a small number of extra rounds is used. However, this behaviour stops as soon as r is greater than $\log_{k+1}(T)$. Indeed, the main cost becomes $2/k$. This can be explained as follows: to have an unit length message broadcast in a whole network under the k -ports mode, a transmission cost of $1/k$ is required to make the message go out from the source and also to make it enter into any other node. We could think that the pipeline techniques could decrease this cost as they should allow to have these two costs covered by each other, in order to go below a total cost of $2/k$. Unfortunately, Proposition 22 proves that no pipelined algorithm can be usefull if $r < T - 1$.

5 Broadcasting with many extra rounds

In this section, we investigate the case where many extra rounds are used, *i.e.* with $r \geq T - 1$. In this range pipelined protocols are possible and we study how such techniques allow to decrease the transmission cost.

5.1 Upper bounds: pipelined algorithms

Proposition 25 For $r \geq T - 1$, $f_T(r) \leq \frac{T+r}{kr+1}$.

Proof. This upper bound follows from an algorithm. In the following \mathbb{Z}_q will denote the set of integers modulo q . The elements of \mathbb{Z}_q will be taken in the set $\{0, 1, \dots, q - 1\}$. The vertex set of the complete graph K_N is represented by \mathbb{Z}_{k+1}^T . The elements in \mathbb{Z}_{k+1}^T may be expressed in the canonical base $\{e_i\}_{1 \leq i \leq T}$, that is, we may denote a vertex $x = (x_1, x_2, \dots, x_T) \in \mathbb{Z}_{k+1}^T$ by $\sum_{i=1}^T x_i \cdot e_i$. The source is the element with all components

equal to 0. In the following, we need to define a *labelling* of the arcs to be used in our algorithm. For each node x , for $1 \leq j \leq k$ and $1 \leq i \leq T$, the arc $(x, x + j \cdot e_i)$ will be labelled by i . In our algorithm, only arcs labelled by i are used during a round $t = i \bmod T$.

Now, let U be the set of vertices such that $U = \{u \in V(K_N) \mid u = j \cdot e_i, \text{ with } 1 \leq j \leq k \text{ and } 1 \leq i \leq T\}$. For each node $u \in U$, let $\mathcal{T}_0(u)$ be simply the vertex u . Now, for $1 \leq h \leq T$, let $\mathcal{T}_h(u)$ be the tree induced from $\mathcal{T}_{h-1}(u)$ by adding all the possible arcs labelled by $(i + h \bmod T)$. Thus, $\mathcal{T}_T(u)$ is a directed spanning tree rooted in u .

In this way, we have defined $k \cdot T$ arc-disjoint spanning trees and the algorithm mainly consists in pipelining concurrently in each of them. Moreover, to maintain the full use of the bandwidth until the end of the process, the last piece of information is broadcast during the last T rounds using a spanning tree rooted in the source. Now, we describe the algorithm.

- The source first cuts equally its unit length message into $kr + 1$ distinct submessages, denoted by m_l with $1 \leq l \leq kr + 1$.
- For $1 \leq t \leq r$, at round t the source sends submessage $m_{(t-1)k+j}$, for $1 \leq j \leq k$, to node $u = j \cdot e_{(t \bmod T)}$.
- When a submessage m_l , with $1 \leq l \leq kr$, has been sent from the source to a node $u \in U$, it is broadcast in $\mathcal{T}_T(u)$ during the next T rounds (using thereby all the arcs labelled by i at each round $t = i \bmod T$).
- The last submessage m_{kr+1} is broadcast from the source in the T last rounds, using all the possible arcs labelled by $t \bmod T$ at round t .

The different rounds use the same transmission cost equal to $1/(kr + 1)$, that is the length of each submessage. This leads to the expected total cost $(T + r)/(kr + 1)$. \square

5.2 Lower bounds

Lemma 26 *For any $r \geq 0$, if $f_T(r) \geq \frac{T+r}{kr+1}$ then $f_T(r + T) \geq \frac{T+(r+T)}{k(r+T)+1}$.*

Proof. Consider an optimal algorithm broadcasting an unit length message in $T + (T + r)$ rounds. After round T , at least $(1 - k\mathcal{L}_T)$ information remains to be broadcast from the source in $T + r$ extra rounds, hence:

$$f_T(r + T) \geq \mathcal{L}_T + f_T(r) \cdot (1 - k\mathcal{L}_T) \geq \mathcal{L}_T + \frac{T + r}{kr + 1} \cdot (1 - k\mathcal{L}_T) = \varphi_1(\mathcal{L}_T)$$

Note also that after time T , an amount of $(N - 1 - \mathcal{I}_T)$ information has to be transmitted on at most $k(N - 1)$ links at each round (as those incoming the source are useless). From Lemma 11, we know that $\mathcal{I}_T \leq \frac{\mathcal{L}_T}{T}(N - 1)$, hence:

$$f_T(T + r) \geq \mathcal{L}_T + \frac{1}{k} \left(1 - \frac{\mathcal{L}_T}{T} \right) = \varphi_2(\mathcal{L}_T)$$

As in the proof of Lemma 11, φ_1 is a decreasing function while φ_2 is an increasing function. So $\min(\varphi_1(\mathcal{L}_T), \varphi_2(\mathcal{L}_T)) \geq \varphi_1(\xi)$, where ξ satisfies $\varphi_1(\xi) = \varphi_2(\xi)$. Solving the equation we get:

$$\begin{aligned} \xi + \frac{T+r}{kr+1}(1 - k\xi) &= \xi + \frac{1}{k}(1 - \frac{\xi}{T}) \\ \xi &= \frac{T}{k(r+T)+1} \\ \varphi_1(\xi) = \varphi_2(\xi) &= \frac{T+(r+T)}{k(r+T)+1} \end{aligned}$$

So we have $f_T(T + r) \geq \frac{T+(r+T)}{k(r+T)+1}$. □

Lemma 27 $f_T(T + 2) = \frac{T+(T+2)}{k(T+2)+1}$

Proof. From Proposition 25 we have $f_T(T + 2) \leq \frac{T+(T+2)}{k(T+2)+1}$ and by using the same arguments we get:

$$\begin{aligned} \frac{T+(T+2)}{k(T+2)+1} \geq f_T(T + 2) &\geq \mathcal{L}_{T+1} + f_T(1) \cdot (1 - k\mathcal{L}_{T+1}) \\ \frac{T+(T+2)}{k(T+2)+1} &\geq \mathcal{L}_{T+1} + \frac{T+1}{k+1}(1 - k\mathcal{L}_{T+1}) \\ \mathcal{L}_{T+1} &\geq \frac{T+1}{k(T+2)+1} \end{aligned}$$

Let us denote by \mathcal{L}'_{T+1} the sum of the transmission costs of the last $T + 1$ rounds, hence $f_T(T + 2) = \mathcal{L}_{T+1} + \mathcal{L}'_{T+1}$.

$$\begin{aligned} \frac{T+(T+2)}{k(T+2)+1} \geq f_T(T + 2) &\geq f_T(1) \cdot (1 - k\mathcal{L}'_{T+1}) + \mathcal{L}'_{T+1} \\ \frac{T+(T+2)}{k(T+2)+1} &\geq \frac{T+1}{k+1}(1 - k\mathcal{L}'_{T+1}) + \mathcal{L}'_{T+1} \\ \mathcal{L}'_{T+1} &\geq \frac{T+1}{k(T+2)+1} \end{aligned}$$

So the lemma holds. □

Theorem 28 For $r \geq T$, and $r = 0, 1, 2 \bmod T$, $f_T(r) = \frac{T+r}{kr+1}$.

Proof. We use Lemma 26 inductively. To initiate the induction for $r = 0 \bmod T$, we use Proposition 8, that is $f_T(0) = T$ verifying the condition of Lemma 26. For $r = 1 \bmod T$, we use Proposition 15, that is $f_T(1) = \frac{T+1}{k+1}$. Finally, for $r = 2 \bmod T$, we use Lemma 27 giving the exact value of $f_T(T+2)$, that verifies the needed condition too. For these three cases, we obtain $f_T(r) \geq \frac{T+r}{kr+1}$. The result follows from Proposition 25. \square

Corollary 29 For $T \in \{1, 2, 3\}$ and for any $r \geq T$, $f_T(r) = \frac{T+r}{kr+1}$.

Corollary 30 For $r \geq T$, $f_T(r) = \frac{T+r}{kr+1} - \mathcal{O}\left(\frac{1}{r^2}\right)$.

Proof. Assume that $r_1 = aT$, $r = aT + b$ and $r_2 = (a+1)T$ with $a \geq 1$ and $0 \leq b < T$. We know by Theorem 28 that $f_T(r_1) = \frac{(a+1)T}{kaT+1}$, and $f_T(r_2) = \frac{(a+2)T}{k(a+1)T+1}$. As by Property 4, f_T is a non-increasing function, we have:

$$\begin{aligned} f_T(r_1) &\geq f_T(r) \geq f_T(r_2) \\ \frac{(a+1)T}{kaT+1} &\geq f_T(r) \geq \frac{(a+2)T}{k(a+1)T+1} \end{aligned}$$

So we obtain for $r \geq T$,

$$\begin{aligned} f_T(r) &= \frac{1}{k} + \frac{1}{ka} - \frac{1}{k^2Ta} - \mathcal{O}\left(\frac{1}{a^2}\right) \\ f_T(r) &= \frac{1}{k} + \frac{T}{kr} - \frac{1}{k^2r} - \mathcal{O}\left(\frac{1}{r^2}\right) = \frac{T+r}{kr+1} - \mathcal{O}\left(\frac{1}{r^2}\right) \end{aligned}$$

\square

The above proposition and corollary make the following conjecture very likely to hold:

Conjecture 31 For any T and any $r \geq T$, $f_T(r) = \frac{T+r}{kr+1}$.

6 Conclusion

In this paper, we have shown the existence of a trade-off between the number of rounds and the transmission cost of broadcast protocols. We have derived tight bounds on the minimum communication time $(T+r) \cdot \alpha + f_T(r) \cdot L \cdot \tau$ of any algorithm using $T+r$ rounds for broadcasting a message of length L in the bidirectional k -ports complete network K_N , with $N = (k+1)^T$. The following table summarizes the different results obtained:

Number of extra rounds	Minimum transmission cost	Comment
$r = 0$	$f_T(0) = T$	greedy
$r = 1$	$f_T(1) = \frac{T+1}{k+1}$	exponential trade-off
$r = 2$	$f_T(2) = \frac{T}{(k+1)^2} + \frac{2}{k+1} - \Theta\left(\frac{1}{kT}\right)$	
$2 \leq r \leq \log_{k+1}(T) + 1$	$f_T(r) = \frac{T}{(k+1)^r} + \Theta\left(\frac{1}{k}\right)$	
$\log_{k+1}(T) + 1 < r < T - 1$	$f_T(r) = \frac{2}{k} + \mathcal{O}\left(\frac{1}{k^2}\right)$	no trade-off
$r \geq T, r = 0, 1, 2 \bmod T$	$f_T(r) = \frac{T+r}{kr+1}$	linear trade-off
$r \geq T - 1$	$f_T(r) = \frac{T+r}{kr+1} - \mathcal{O}\left(\frac{1}{r^2}\right)$	

Our lower bounds have been obtained for complete networks, therefore they are valid for any topology. Conversely, our matching upper bounds derived from our algorithms require strong connectivity to be reached. For example, we can show that the exponential decrease of the transmission cost does not occur in a ring network. However, all routing mechanisms based on the circuit-switched mode can take advantage of our algorithms, due to the additional connectivity provided. In particular, the *Wavelength Division Multiplexed* (WDM) optical routing can offer full connectivity in various network topologies (see [3, 2]). Our study has thus been fruitfully taken into account in [1] for multi-hop optical ring and mesh networks. Further work should concentrate on other less connected topologies, for which greater lower bounds will arise.

Acknowledgements

We are grateful to J-C. Bermond for his helpful remarks and comments.

References

- [1] B. Beauquier. Broadcasting in WDM optical rings and tori. In *Proc. the DIMACS Workshop : "Multichannel Optical Networks : Theory and Practice"*, volume 46 of *DIMACS Series on Discrete Mathematics and Theoretical Computer Science*, pages 63–73, Rutgers University, NJ, March 1998. American Mathematical Society.
- [2] B. Beauquier, J-C. Bermond, L. Gargano, P. Hell, S. Pérennes, and U. Vaccaro. Graph problems arising from wavelength-routing in all-optical networks. In *Proc. of 2nd IEEE Workshop on Optics and Computer Science (WOCS, part of IPPS'97)*, Geneva, Switzerland, April 1997. IEEE Press.
- [3] J-C. Bermond, L. Gargano, S. Pérennes, A. A. Rescigno, and U. Vaccaro. Efficient collective communication in optical networks. In *Proc. of 23rd International Colloquium on Automata, Languages and Programming (ICALP'96)*, volume 1099 of *Lecture Notes in Computer Science*, pages 574–585. Springer-Verlag, July 1996.
- [4] J-C. Bermond, P. Michallon, and D. Trystram. Broadcasting in wraparound meshes with parallel monodirectional links. *Parallel Computing*, 18:639–648, 1992.
- [5] J. Bruck, C-T. Ho, S. Kipnis, and D. Weathersby. Efficient algorithms for all-to-all communications in multi-port message-passing systems. In *Annual ACM Symposium on Parallel Algorithms and Architectures (SPAA '94)*, pages 298–309, Cape May, New Jersey, June 1994. ACM.
- [6] O. Delmas. *Communications par commutation de circuits dans les réseaux d'interconnexion*. Thèse de doctorat, Université de Nice - Sophia Antipolis, U.F.R. de Sciences, Ecole Doctorale - Sciences pour l'Ingénieur, France, January 1997.
- [7] E. Fleury. *Communications, routage et architectures des machines à mémoire distribuée - Autour du routage wormhole*. Thèse de doctorat, Université de Lyon, Ecole Normale Supérieure de Lyon, France, 1996.
- [8] P. Fraigniaud. Performance analysis of broadcasting hypercubes with restricted communication capabilities. *Journal of Parallel and Distributed Computing*, 16:15–26, 1992.
- [9] P. Fraigniaud and C. Laforest. Disjoint spanning trees of small depth. In G.R. Joubert, D. Trystram, F.J. Peters, and D.J. Evans, editors, *ParCo '93*, pages 105–112, 1993.
- [10] P. Fraigniaud and E. Lazard. Methods and problems of communication in usual networks. *Discrete Applied Mathematics*, 53:79–133, 1994. Special volume proceedings international workshop on broadcasting and gossiping 1990.
- [11] S.M. Hedetniemi, S.T. Hedetniemi, and A.L. Liestman. A survey of gossiping and broadcasting in communication networks. *Networks*, 18:319–349, 1986.

- [12] C.T. Ho and S.L. Johnsson. Optimum broadcasting and personalized communication in hypercubes. *IEEE Transactions on Computers*, 38(9):1249–1268, 1989.
- [13] J. Hromkovic, R. Klasing, B. Monien, and R. Peine. *Combinatorial Network Theory*, chapter Dissemination of information in interconnection networks (broadcasting and gossiping), pages 125–212. Kluwer Academic, 1995.
- [14] Jean de Rumeur. *Communication dans les réseaux de processeurs*. Collection Etudes et Recherches en Informatique. Masson, Paris, 1994.
- [15] Q. F. Stout and B. Wagar. Intensive hypercube communication, prearranged communication in link-bound machines. *Journal of Parallel and Distributed Computing*, 10:167–181, 1990.

Contents

1	Introduction	3
2	Framework	4
3	Upper bound for broadcasting with a few extra rounds	6
4	Lower bounds for broadcasting with a few extra rounds	8
4.1	Notation	8
4.2	Preliminaries	8
4.3	One extra round	10
4.4	Two extra rounds	11
4.5	Some extra rounds	13
5	Broadcasting with many extra rounds	14
5.1	Upper bounds: pipelined algorithms	14
5.2	Lower bounds	15
6	Conclusion	17



Unité de recherche INRIA Sophia Antipolis
2004, route des Lucioles - B.P. 93 - 06902 Sophia Antipolis Cedex (France)

Unité de recherche INRIA Lorraine : Technopôle de Nancy-Brabois - Campus scientifique
615, rue du Jardin Botanique - B.P. 101 - 54602 Villers lès Nancy Cedex (France)

Unité de recherche INRIA Rennes : IRISA, Campus universitaire de Beaulieu - 35042 Rennes Cedex (France)

Unité de recherche INRIA Rhône-Alpes : 655, avenue de l'Europe - 38330 Montbonnot St Martin (France)

Unité de recherche INRIA Rocquencourt : Domaine de Voluceau - Rocquencourt - B.P. 105 - 78153 Le Chesnay Cedex (France)

Éditeur
INRIA - Domaine de Voluceau - Rocquencourt, B.P. 105 - 78153 Le Chesnay Cedex (France)
<http://www.inria.fr>
ISSN 0249-6399